nature communications

Article

https://doi.org/10.1038/s41467-025-57332-5

Blood and tissue HIV-1 reservoirs display plasticity and lack of compartmentalization in virally suppressed people

Received: 28 August 2024

Accepted: 14 February 2025

Published online: 04 March 2025

Check for updates

Marion Pardons[1,6], Laurens Lambrechts[1,2,6], Ytse Noppe[1], Liesbet Termote[1], Sofie De Braekeleer[1], Jerel Vega[3], Ellen Van Gulck[4], Sarah Gerlo[1,5] & Linos Vandekerckhove[1]

Characterizing the HIV-1 reservoir in blood and tissues is crucial for the development of curative strategies. Using an HIV Tat mRNA-containing lipid nanoparticle (Tat-LNP) in combination with panobinostat, we show that p24+ cells from blood and lymph nodes exhibit distinct phenotypes. Blood p24+ cells are found in both central/transitional (TCM/TTM) and effector memory subsets, mostly lack CXCR5 expression and are enriched in GZMA+ cells. In contrast, most lymph node p24+ cells display a TCM/TTM phenotype, with approximately 50% expressing CXCR5 and nearly all lacking GZMA expression. Furthermore, germinal center T follicular helper cells do not appear to harbor the translation-competent reservoir in long-term suppressed individuals. Near full-length HIV-1 sequencing in longitudinal samples from matched blood, lymph nodes, and gut indicates that clones of infected cells, including those carrying an inducible provirus, persist and spread across various anatomical compartments. Finally, uniform genetic diversity across sites suggests the absence of ongoing replication in tissues under treatment.

Because most HIV-1 studies have focused on the peripheral blood, HIV-1 persistence in tissues remains poorly documented. However, tissues serve as ideal sanctuaries for HIV reservoir persistence. Indeed, some reports suggest suboptimal penetration of antiretroviral drugs in anatomical sites[1,2], potentially allowing for low levels of ongoing replication. Furthermore, some anatomical compartments are defined by a naturally reduced susceptibility to immune responses, such as the germinal centers in lymph nodes from which the cytotoxic CD8 T cells are excluded[3]. A non-human primate study on SIV-infected macaques under antiretroviral therapy (ART) has revealed that the majority of the total body reservoir is found in the gut (> 98%), though SIV-infected cells were detected in every organ system examined[4]. Similar results were obtained in human autopsy studies[5–7], where HIV-infected cells were detected in almost all tissues, with the lymph nodes and gastrointestinal tissues representing hotspots for the persistence of intact proviruses[7]. In line with this observation, a recent study in SIV-infected macaques has shown that lymphoid tissues may represent a preferential source of viral rebound following treatment interruption[8].

Characterizing the HIV-1 reservoir not only in blood but also in tissues is key to develop cure interventions. The fast development of single-cell technologies has greatly facilitated the phenotypic characterization of the HIV-1 reservoir[9–15]. However, characterizing the translation-competent reservoir ex vivo has been hampered by the extremely low frequencies of cells producing viral proteins during viral suppression. While strong mitogens such as phorbol myristate acetate (PMA) and anti-CD3/CD28 antibodies can be employed to force

[1]HIV Cure Research Center, Department of Internal Medicine and Pediatrics, Ghent University Hospital, Ghent University, Ghent, Belgium. [2]BioBix, Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium. [3]Arcturus Therapeutics, 10628 Science Center Drive, Suite 250, San Diego, California, USA. [4]Johnson & Johnson Innovative Medicine, Janssen Pharmaceutica NV, Beerse, Belgium. [5]Department of Biomolecular Medicine, Ghent University, Ghent, Belgium. [6]These authors contributed equally: Marion Pardons, Laurens Lambrechts.
e-mail: linos.vandekerckhove@ugent.be

reactivation from latency and boost viral production, these molecules lack specificity[16] and severely modify the phenotype of CD4 T cells[17]. Reactivation of the viral reservoir using potent latency reversing agents (LRAs) that do not alter the cellular phenotype, such as the recently described Tat mRNA-containing nanoparticle Tat-LNP[18], would greatly facilitate the detection and phenotypic characterization of the inducible viral reservoir in blood and tissues.

The parallel development of sequencing technologies has also expanded the capability to study the viral composition of the HIV-1 reservoir and its genetic diversity. Several studies have attempted to compare the HIV-1 proviral landscape between blood and tissues[19–27]. Importantly, most of these studies sequenced sub-genomic regions rather than full-length proviruses, thereby increasing the risk of identifying inaccurate links between HIV sequences[28], as these proviruses can differ in the rest of their genome. Moreover, the obtained sub-genomic datasets lack information about proviral intactness, a feature that remains understudied within tissues. While full-length proviral sequencing has been recently conducted in two human post-mortem studies[6,7], the absence of blood collection before death impedes the ability to compare the viral composition of the reservoir between blood and tissues.

In this study, we aimed to address several key questions: (i) Do translation-competent reservoir cells in matched blood and lymph nodes exhibit distinct phenotypic features in ART-treated individuals? (ii) Are these reservoirs enriched in germinal center T follicular helper (GC Tfh) cells in individuals undergoing treatment for more than 10 years? (iii) Does near full-length sequencing reveal compartmentalization between blood and tissue samples? (iv) Do infected cell clones display distinct phenotypes based on their anatomical location? To explore these questions, we simultaneously characterized the phenotype of the translation-competent reservoir and the HIV-1 proviral landscape in matched blood, lymph nodes, and gut biopsies from ART-suppressed individuals. We previously showed that Tat-LNP/panobinostat (PNB) induces latency reversal in a significantly higher proportion of latently infected cells compared to PMA/ionomycin (≈ 4-fold higher)[18]. Therefore, we took advantage of this unique property to investigate the translation-competent reservoir in matched blood and lymph nodes using the flow cytometry-based HIV-Flow assay[11]. As Tfh cells were previously identified as preferential HIV reservoirs during ART[29–31], we included markers defining this population in our analyzes (CXCR5, PD1, ICOS). We also assessed the expression of two intracellular markers (GZMA and Bcl2) that were previously identified by our team and others to be preferentially expressed by HIV reservoir cells[18,32–34]. Finally, the recently described HIV Proviral UMI-mediated Long-read Sequencing (HIV-PULSE) assay[35] was employed to comprehensively assess the viral composition of the HIV-1 reservoir in matched blood and tissue samples.

Results

Experimental approach to study the phenotype and the proviral composition of the HIV-1 reservoir in matched blood, lymph nodes and gut

This study included nine participants on suppressive ART, four provided matched blood/lymph node/gut biopsies, and five contributed samples from two anatomical compartments (T1; Fig. 1, Supplementary Table 1). For six out of nine participants, a blood sample from a previous collection time point was available (T0), allowing for longitudinal assessment of the reservoir dynamics. The duration between these two time points ranged from 0.8 to 2 years. All participants but one were treated for more than 10 years (median = 14.6 years). To ensure the detection of p24+ cells in both blood and lymph nodes at T1, participants were included in this study if their frequency of p24+ cells in the blood at T0 exceeded 10 per 10^6 CD4 T cells following Tat-LNP/PNB stimulation.

The proviral composition of the total and inducible reservoirs from blood samples at T0 was assessed using the HIV-PULSE[35] and STIP-Seq[36] assays, respectively (Fig. 1). Isolated CD4 T cells from both blood and lymph nodes at T1 were cultured for 24h in presence of Tat-

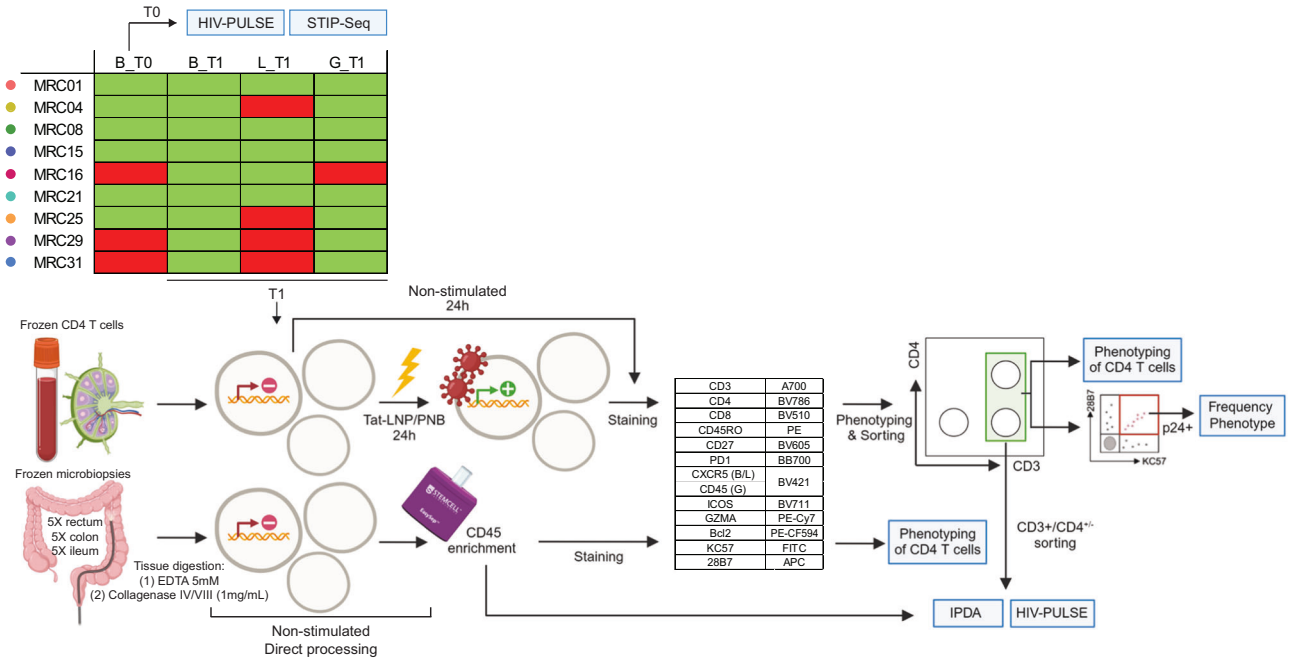


Fig. 1 | Experimental approach overview. HIV-PULSE (total reservoir) and STIP-Seq (inducible reservoir) sequences were obtained from blood samples at T0. CD4 T cells from matched peripheral blood and lymph nodes at T1 were stimulated for 24h with Tat-LNP/PNB. Following staining with the specified panel of antibodies, immunophenotyping of CD4 T cells and p24+ cells was documented using the cell sorter while simultaneously sorting live $CD3^+CD4^{+/-}$ T cells for downstream virological assays. Magnetic cell isolation was used to enrich for $CD45^{+/-}$ cells from 15 gut biopsies. A portion of these cells was stained with the indicated antibody panel, while the remainder was lysed for downstream virological assays. B = blood, L = lymph node, G = gut. Created in BioRender. Pardons, M. (2025) https://BioRender.com/u36d150.

LNP/PNB (Fig. 1). To ensure an effective comparison of HIV reactivation between blood and lymph nodes, cells from both compartments were cultured in parallel for each participant. To assess the impact of Tat-LNP/PNB on the cell phenotype, a small fraction of the cells remained unstimulated. CD4 T cells were stained for cell surface markers, fixed/permeabilized and subjected to the HIV-Flow protocol[11] to assess the frequency and phenotype of p24-producing cells. To perform immunological and virological assays on the same cell cultures, immunophenotyping was documented using the cell sorter while simultaneously sorting live $CD3^+CD4^{+/-}$ T cells for downstream virological assays. Qualitative reservoir assessment was provided by the HIV-PULSE assay[35], while quantitative assessment was achieved using a multiplex digital PCR-based version of the Intact Proviral DNA Assay (IPDA)[37]. Of note, in the case of lymph node cells, because the stimulation was performed on a lower number of CD4 T cells compared to blood samples, the dead $CD3^+CD4^{+/-}$ fraction was also sorted to increase the number of HIV sequences retrieved for each participant. For gut biopsies, we utilized magnetic cell isolation to enrich for $CD45^{+/-}EpCAM^-$ cells (Supplementary Fig. 1). Because of the limited number of gut immune cells obtained from 15 biopsies and their fragility during culture, Tat-LNP/PNB stimulation was not performed. Following enrichment, a small fraction of $CD45^{+/-}$ cells was directly stained with the same panel of antibodies used for blood and lymph nodes, except for CXCR5 due to its loss of expression following collagenase digestion. The remaining gut cells were lysed for subsequent virological assays[35,37].

Frequencies of p24+ cells are similar between matched blood and lymph nodes following Tat-LNP/PNB stimulation

We first assessed the viability of CD4 T cells from blood and matched lymph nodes following Tat-LNP/PNB stimulation, using Live/Dead staining (Fig. 2A). As previously reported[18], incubation of peripheral CD4 T cells with Tat-LNP/PNB had a limited impact on their viability (means = 87.4% and 76.6% in non-stimulated vs Tat-LNP/PNB-stimulated cells, respectively). However, Tat-LNP/PNB treatment exhibited a more profound impact on the viability of lymph node CD4 T cells (means = 77.1% and 53.7% in non-stimulated vs Tat-LNP/PNB-stimulated cells, respectively). This may be due to the mechanical dissociation process, which could make lymph node cells more fragile and therefore more susceptible to LRA treatment. Using the HIV-Flow assay, we measured the frequency of p24+ cells in both anatomical sites (Fig. 2B). While frequencies of p24+ cells were higher in the lymph nodes compared to the blood in 2/5 participants, the opposite trend was observed in the remaining individuals (means = 11.5 and 9.0 p24+ cells/10^6 CD4 T cells in blood vs lymph nodes). Furthermore, the mean fraction of intact proviruses (intact/total HIV), as measured by IPDA in the sorted live $CD3^+CD4^{+/-}$ fraction, was 10.5% in the blood and 8.5% in lymph nodes, with notable intra- and inter-individual variability (Fig. 2C).

P24+ cells from blood and lymph nodes display distinct phenotypes

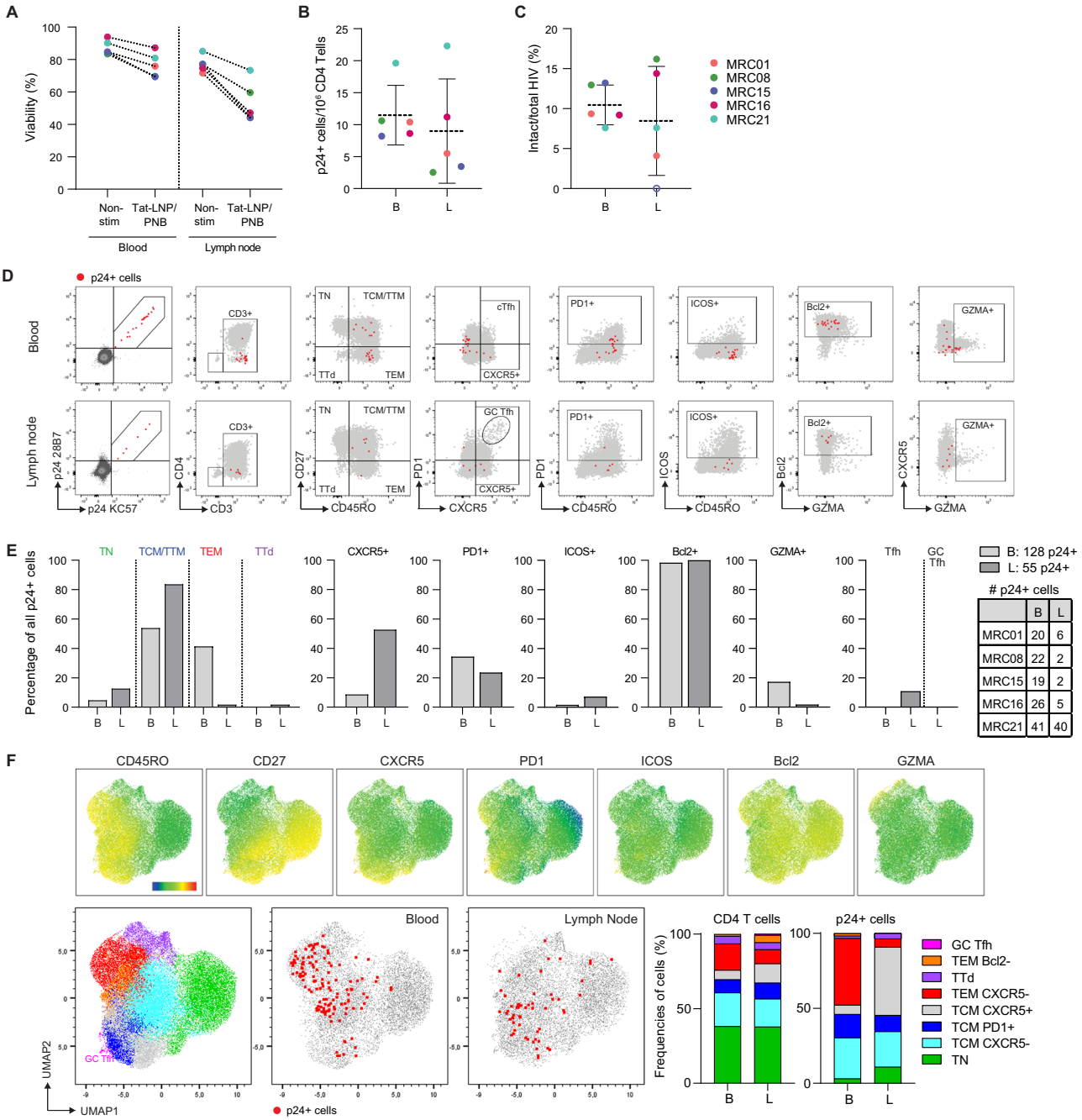
We first compared the ex vivo phenotype of CD4 T cells across matched blood and tissues. Manual gating (Supplementary Figs. 2A–B) and unsupervised clustering through FlowSOM[38] analysis (Supplementary Fig. 2C) confirmed differences in subset composition across anatomical compartments: (i) Naïve CD4 T cells (TN) were virtually absent from the gut (means = 32.3%, 39.1%, 2.6% in blood, lymph nodes and gut, respectively), (ii) as expected, CXCR5 was expressed at higher levels in the lymph nodes compared to the blood (means = 16.1% and 34.0% in blood and lymph nodes, respectively), (iii) the vast majority of CD4 T cells expressed Bcl2 in all anatomical sites (means = 93.8%, 93.7%, 89.0% in blood, lymph nodes and gut, respectively), (iv) frequencies of GZMA+ cells displayed high variability between participants in the blood CD4 T cells (range = 1.4%-45.2%, mean = 14.8%), whereas these frequencies dropped markedly in the tissues (means = 1.7% and 5.1% in the lymph nodes and gut, respectively). The frequency of GC Tfh cells ($CXCR5^{high}PD1^{high}$) in CD4 T cells from the lymph nodes ranged from 0.2 to 4.6% (mean = 1.3%) (Supplementary Fig. 2B).

We also determined the impact of Tat-LNP/PNB stimulation on the expression of the markers included in this study (Supplementary Fig. 3A). Frequencies of CD4 T cell subsets, CXCR5+, PD1+ and Tfh cells ($CXCR5^{int}PD1^{int}$) were comparable between non-stimulated and treated cells in both anatomical compartments. However, the frequency of GC Tfh cells ($CXCR5^{high}PD1^{high}$) in the lymph nodes decreased in the participant with the highest frequency of this subset (MRC01: 4.4% and 1.1% GC Tfh cells in non-stimulated vs Tat-LNP/PNB-treated cells, respectively). Additionally, the frequencies of GZMA+ cells decreased in all participants in both compartments following stimulation (means in blood = 15.7% and 7.4%, means in lymph nodes = 2.1% and 1.1% in non-stimulated vs Tat-LNP/PNB-stimulated cells, respectively). Despite the reduced expression of GZMA in bulk CD4 T cells following Tat-LNP/PNB treatment, blood-derived p24+ cells still displayed preferential enrichment in GZMA+ cells in all five participants (Supplementary Fig. 3B), further confirming our previously reported findings using Tat-LNP alone[18]. In contrast, blood-derived p24+ cells were not preferentially observed in the CXCR5+, ICOS+, Bcl2+ subsets, nor in the circulating Tfh cells ($CXCR5^{int}PD1^{int}$) (Supplementary Fig. 3B). Of note, blood-derived p24+ cells did not exhibit a higher mean fluorescence intensity for Bcl2 compared to p24- cells (mean MFIs = 4585 vs 4200 in p24- vs p24+ cells) (Supplementary Fig. 3C).

We then compared the phenotype of p24+ cells in paired blood and lymph nodes by manual gating (Fig. 2D, E). Due to the limited number of p24+ cells obtained in the lymph nodes for 4/5 participants (2-6 p24+ cells), all lymph node-derived p24+ cells from the five ART-suppressed individuals were pooled ($n = 55$ p24+ cells) and their phenotype was compared to the aggregated p24+ cells from blood ($n = 128$ p24+ cells) (Fig. 2E). While peripheral p24+ cells were predominantly found in the TCM/TTM and TEM compartments, lymph node-derived p24+ cells were mainly present in the TCM/TTM subset and virtually absent from TEM cells. Similar proportions of p24+ cells expressed PD1 in blood and lymph nodes (44/128: 34.4% and 13/55: 23.6%, respectively). Only 11/128 (8.6%) peripheral p24+ cells expressed the follicular homing marker CXCR5, whereas 29/55 (52.7%) p24+ cells from the lymph nodes expressed CXCR5. However, in both compartments, few p24+ cells displayed a Tfh phenotype ($CXCR5^{int}PD1^{int}$: 0/128 in blood and 6/55 p24+ cells in lymph nodes). Interestingly, none of the lymph node-derived p24+ cells exhibited characteristics of GC Tfh cells ($CXCR5^{high}PD1^{high}$), suggesting that this cellular subset may not harbor the translation-competent reservoir in long-term treated individuals. Moreover, almost all p24+ cells from the blood and lymph nodes expressed Bcl2, in line with the majority of CD4 T cells being Bcl2+ in these anatomical sites (Supplementary Fig. 2B). Finally, while 22/128 (17.2%) peripheral p24+ cells expressed the cytotoxic molecule GZMA, only 1/55 (1.8%) p24+ cells was GZMA+ in the lymph nodes. This observation indicates that, in contrast with the blood, GZMA expression does not appear to be a distinctive feature of the inducible reservoir in lymph nodes.

Finally, we used the unsupervised clustering provided by FlowSOM[38] to compare the phenotype of CD4 T cells and p24+ cells between blood and lymph nodes (Fig. 2F, Supplementary Fig. 3D). Eight clusters were defined using this approach. This analysis confirmed that p24+ cells from the blood and the lymph nodes clustered in different subsets of cells. While 44.5% of the p24+ cells clustered in the TEM CXCR5- cluster in the blood, only 5.5% of the p24+ cells displayed this phenotype in the lymph nodes. In contrast, 45.5% of the p24+ cells displayed a TCM CXCR5+ phenotype in the lymph nodes, whereas only 6.3% of the p24+ cells clustered in this subset in the blood.

https://doi.org/10.1038/s41467-025-57332-5



	B	L
MRC01	20	6
MRC08	22	2
MRC15	19	2
MRC16	26	5
MRC21	41	40

Fig. 2 | Immunophenotyping of p24+ cells in matched blood and lymph nodes. CD4 T cells from $n = 5$ ART-suppressed individuals were cultured for 24h in the absence of stimulation or in the presence of Tat-LNP/PNB. **A** Percentages of live cells (as defined by a negative Live/Dead stain) among all recorded events. **B** Frequencies of p24+ cells measured by the HIV-Flow assay following stimulation with Tat-LNP/PNB. **C** Intact/total HIV ratios (%) in sorted $CD3^+CD4^{+/-}$ T cells, as defined by IPDA. **D** Representative flow cytometry dot plots showing the expression of the markers of interest in total CD4 T cells (gray dots) and in p24+ cells (red dots) in blood (upper panel) and lymph nodes (lower panel) (MRC01). Supplementary Fig. 3A displays the frequencies of cell subsets for each participant. **E** Percentage of p24+ cells from blood (B) and lymph nodes (L) with a given phenotype. All p24+ cells from each participant were pooled together to perform this analysis. **F** Uniform Manifold Approximation and Projection (UMAPs) derived from the FlowSOM analysis (8 clusters) and combining Tat-LNP/PNB-stimulated $CD3^+CD4^{+/-}$ T cells from blood and lymph nodes. The expression levels of each marker included in this analysis are shown on the top panel. P24+ cells are marked by red dots and are overlaid onto the UMAPs of CD4 T cells from blood and lymph nodes. The bar plots show the percentage of each cluster within CD4 T cells and p24+ cells. TN = naïve T cells, TCM/TTM = central/transitional memory T cells, TEM = effector memory T cells, TTd = terminally differentiated T cells. Means and standard deviations are plotted on Fig. 2B, C. Statistics were not performed for Fig. 2A–C due to the limited number of participants ($n = 5$). Source data are provided as a Source Data file.

Assessment of the reservoir's proviral composition in matched blood and tissues

To align our phenotypic observations with virological data on the reservoir's composition, we utilized the HIV-PULSE assay[35] on blood (T0 and T1) and tissue-derived cells (T1). In total, we obtained 1442 sequences in the blood T0 (1023 distinct), 1011 in the blood T1 (734 distinct), 709 in the lymph nodes T1 (626 distinct) and 259 in the gut T1 (228 distinct). The numbers of distinct and total proviral sequences per participant are indicated in the table, with an average of 380 (263 distinct) sequences per participant (Fig. 3A). Differences

https://doi.org/10.1038/s41467-025-57332-5

A

Participant	Distinct	Total
MRC01	220	445
MRC04	355	447
MRC08	627	885
MRC15	62	140
MRC16	66	71
MRC21	687	946
MRC25	210	311
MRC29	65	86
MRC31	79	90
Total	*2371*	*3421*

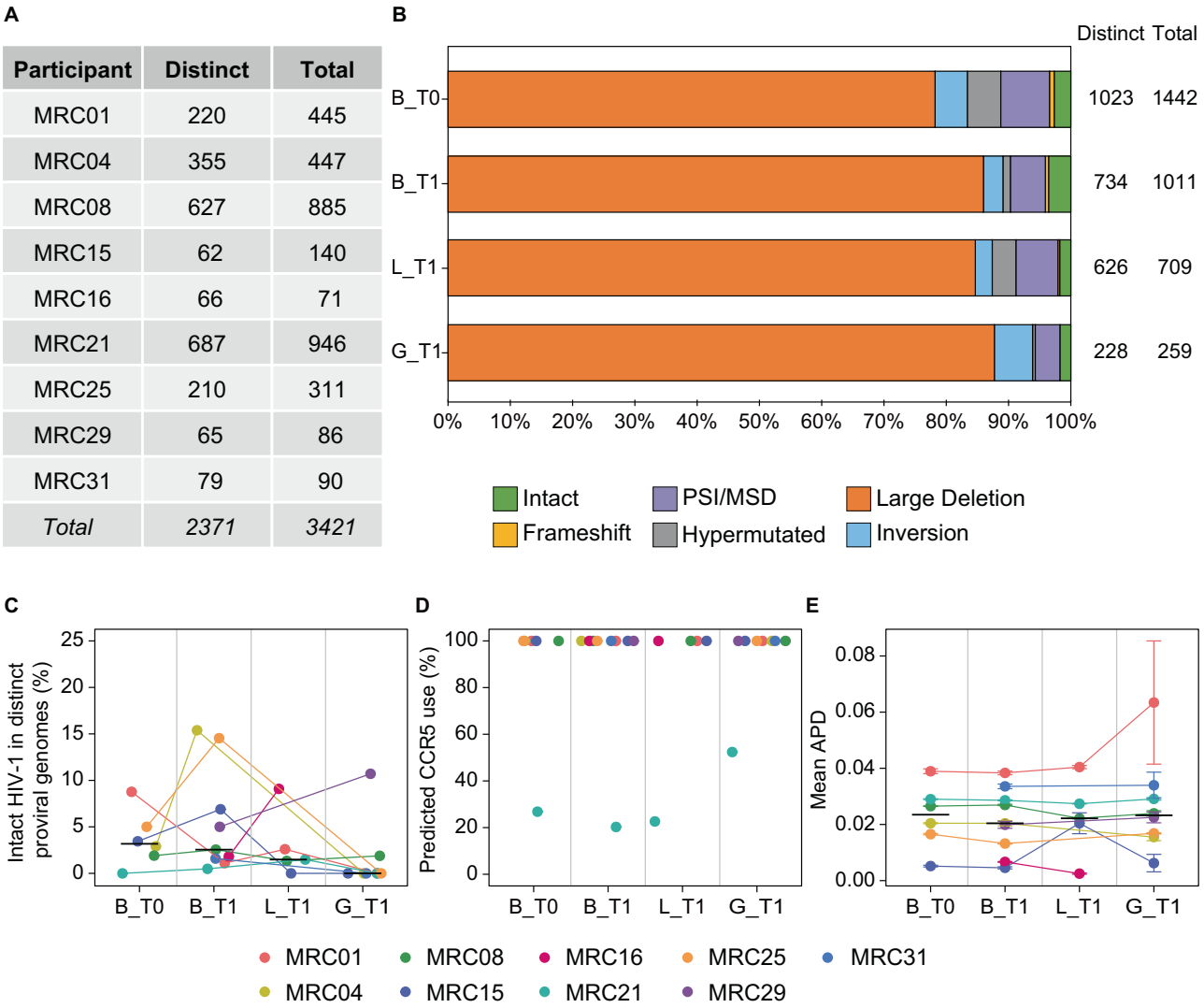


Fig. 3 | Classification, tropism and diversity among HIV-1 proviruses from blood and matched tissues. The proviral reservoir was assayed by the HIV-PULSE assay. **A** The table shows the number of distinct and total proviruses retrieved by HIV-PULSE for each participant. **B** Proportions of different proviral classes observed among the distinct proviruses for each anatomical site (B = blood, L = lymph node, G = gut) and time point (T0 and T1). The numbers of distinct and total proviruses are displayed on the right. **C**–**E** For each participant, within each anatomical site and time point ($n = 6$ participants for B_T0, 9 for B_T1, 5 for L_T1, 8 for G_T1): (**C**) proportions of intact proviruses (for statistics, two-sided paired Wilcoxon signed-rank tests were performed, Supplementary Data 1), (**D**) predicted CCR5 use, and (**E**) average pairwise distance (APD; for statistics, two-sided parametric paired t-tests were performed, Supplementary Data 1. Error bars represent standard deviations). No statistically significant differences were observed in any of the pairwise comparisons. Medians are reported. Source data are provided as a Source Data file.

between distinct and total numbers are due to the repeated detection of identical proviral sequences, including both intact and defective proviruses, and can be attributed to the clonal expansion of HIV-infected cells. Proviral classification revealed that large deletions were the most prevalent defect in all anatomical compartments (78.2%, 86.0%, 84.7%, 87.7% in blood T0, blood T1, lymph nodes and gut, respectively) (Fig. 3B, Supplementary Fig. 4). The proportions of intact HIV-1 genomes showed intra-individual variability between the two time points in the blood (Fig. 3C), likely resulting from several factors such as sampling depth, and/or the waxing and waning of infected clones carrying an intact provirus over time. On average, similar proportions of intact proviruses were detected between time points and anatomical sites (medians = 3.2%, 2.6%, 1.5% and 0.0%, in blood T0, blood T1, lymph nodes and gut, respectively) (Fig. 3C; see Supplementary Data 1 for statistical analyzes). For all participants but one (MRC21), all proviruses recovered by HIV-PULSE were predicted to be CCR5-tropic based on their V3 sequence (Fig. 3D). Interestingly, in MRC21, 14.6% of the blood-derived p24+ cells and 17.5% of the lymph node-derived p24+ cells displayed a TN phenotype (Supplementary Fig. 3B), which aligns with the detection of CXCR4-tropic sequences in that participant. Furthermore, the increased prevalence of CCR5-tropic sequences in the gut of this individual compared to the blood and lymph nodes correlates with the sparse presence of TN cells in the gut (Supplementary Fig. 2). Finally, we analyzed the genetic diversity within blood and tissues-derived sequences, based on alignments of the *nef* gene (Fig. 3E). The average pairwise distance (APD) was not significantly different between both blood time points and tissues (Supplementary Data 1). In addition, to further investigate whether the obtained datasets provided evidence of compartmentalization across tissues and time points, we conducted a series of tests for panmixia. Consistent with the similar APD observed between anatomical sites, both distance and tree-based methods applied to *nef*-based alignments showed no significant signs of compartmentalization (Supplementary Table 2), further supporting the notion that ongoing viral replication leading to distinct HIV populations does not occur in blood and tissues from ART-suppressed individuals[20,26,27,39].

https://doi.org/10.1038/s41467-025-57332-5

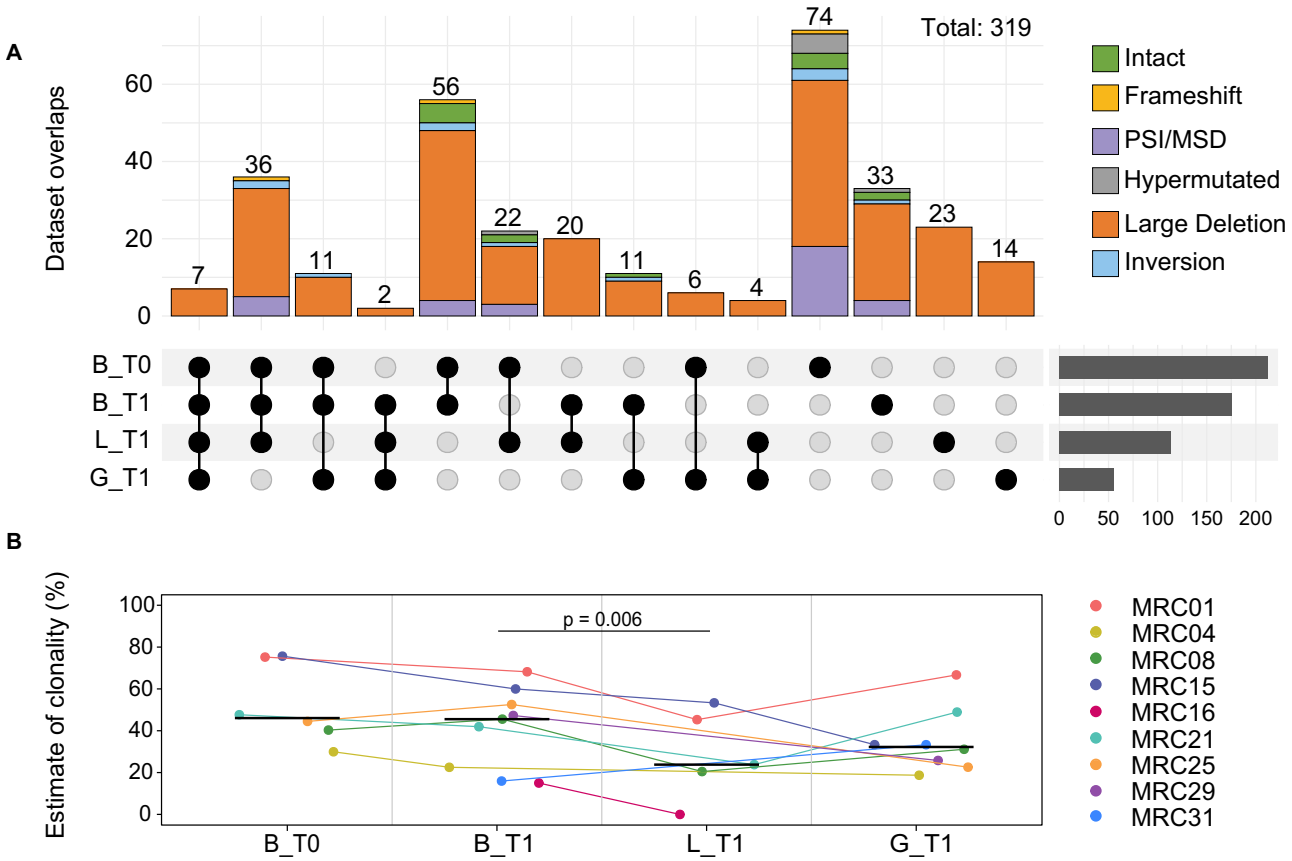


Fig. 4 | Clones of infected cells overlap between time points and anatomical compartments. The proviral reservoir was assayed by the HIV-PULSE assay. **A** UpSet-plot visualizing the number of distinct HIV-PULSE sequences that overlap between anatomical compartments (B = blood, L = lymph node, G = gut) and time points (T0 and T1). Proviral classification is color-coded. The total number of distinct overlapping sequences is indicated on the top of each bar plot. The axis of the bar graph on the right indicates the total number of distinct sequences per dataset. Of note, not all participants contributed to every time point and anatomical compartment (Supplementary Fig. 5, Supplementary Table 3). **B** Estimate of clonality for each participant within each anatomical site and time point ($n = 6$ participants for B_T0, 9 for B_T1, 5 for L_T1, 8 for G_T1), defined as the number of clonal sequences divided by the sum of clonal and non-clonal sequences. Medians are reported. For statistics, two-sided parametric paired t-tests were performed (Supplementary Data 1). Significant p-values are shown. Source data are provided as a Source Data file.

Clones of infected cells overlap between time points and anatomical compartments

Pooling proviral sequences from all participants, we found clonal populations of infected cells that were exclusively retrieved in one compartment, with no overlap with any other time point or anatomical site ($n = 74$ in the blood T0, 33 in the blood T1, 23 in the lymph nodes T1, and 14 in the gut T1) (Fig. 4A, Supplementary Fig. 5). A total of 56 sequences were exclusively shared between blood T0 and T1 and did not overlap with other anatomical sites, five of them being intact. In contrast, we observed 7 distinct sequences that overlapped between all time points and anatomical sites, all of them displaying large internal deletions. In addition, a total of 36 distinct sequences were shared between blood T0/T1 and lymph nodes T1 but not with the gut T1, none of these sequences being intact (77.8% with large internal deletions, 13.9% with PSI/MSD defects, 5.6% with inversions and 2.8% with frameshifts). Furthermore, we observed a total of 11 distinct sequences that overlapped between blood T0/T1 and the gut T1 but not with the lymph nodes T1, none of these sequences being intact (90.9% with large internal deletions and 9.1% with inversions). Overlaps of clonally infected populations carrying intact proviruses were exclusively observed between blood T0 and T1 (8.9%), blood T0 and lymph nodes T1 (9.1%), as well as blood T1 and gut T1 (9.1%). The identification of clonally infected cells across various time points and anatomical compartments confirms the previously reported observation that certain clones can persist for several years and indicates that these clones disseminate into various tissues using the blood as a conduit. Finally, we made an estimation of clonality, defined as the number of clonal sequences divided by the sum of clonal and non-clonal sequences[40]. The clonality estimates highly varied between participants, ranging from 30% to 76% in the blood T0 (Fig. 4B). Clonality estimates were not significantly different between the two time points in the blood (medians = 46% for both time points, $p = 0.347$) (Supplementary Data 1). Interestingly, the clonality estimate was consistently lower in the lymph nodes compared to the blood for each participant (medians = 24% for lymph nodes T1 and 46% for blood T1; $p = 0.006$) (Supplementary Data 1).

Clones of infected cells carrying an inducible provirus overlap between time points and anatomical compartments

We performed phylogenetic analyzes using near full-length HIV sequences (> 7.9 kb) while excluding irrelevant sequences (viruses classified with inversions or hypermutation). These revealed that HIV-PULSE sequences from different anatomical sites do not form distinct phylogenetic clusters but are intermixed within each participant (Fig. 5). To determine if HIV-infected clones recovered from blood and tissues carry inducible proviruses, we also included STIP-Seq sequences, which correspond to HIV proviruses recovered from blood T0-derived p24+ cells following PMA/ionomycin, Tat-LNP/PNB or Tat-LNP/PMA stimulation (Supplementary Data 2). We identified a total of 25 matches between STIP-Seq sequences (retrieved in blood T0) and

https://doi.org/10.1038/s41467-025-57332-5

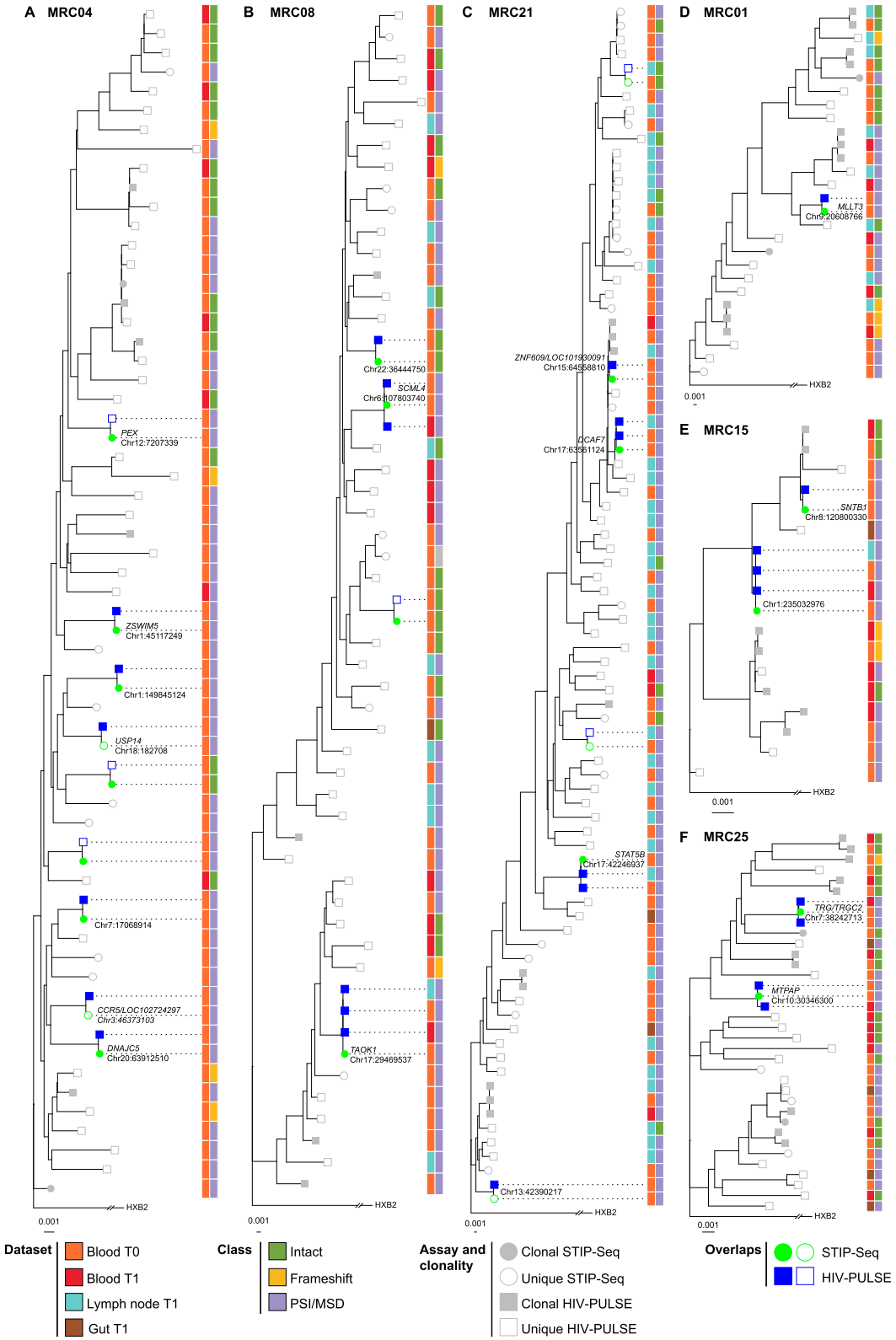


Fig. 5 | Clones of infected cells carrying an inducible provirus overlap between time points and anatomical compartments. A–F Phylogenetic trees including all near full-length distinct proviruses (> 7.9 kb) obtained with HIV-PULSE (blood and tissues at T0/T1) and with STIP-Seq (blood at T0) in $n = 6$ ART-treated participants. Sequences with inversions, large deletions and hypermutations were excluded. Symbols reflect the different assays. Proviruses recovered in a single assay are shown in gray while proviruses overlapping between assays are shown in green (STIP-Seq) or blue (HIV-PULSE). Empty symbols indicate sequences that were retrieved only once (unique) in that respective assay, while full symbols indicate sequences that were retrieved multiple times (clonal). For each provirus, the anatomical site, time point and the provirus classification are indicated. For overlapping sequences between the HIV-PULSE and the STIP-Seq assays, the integration site and chromosomal location are reported when available.

HIV-PULSE sequences (retrieved in blood T0/T1 and tissues T1) (Fig. 5, Supplementary Fig. 6). Of note, among the 25 matches: (i) 15 matches were observed between STIP-Seq blood T0 sequences and HIV-PULSE blood T0 sequences (3 intact and 12 PSI/MSD-defective), but did not overlap with HIV-PULSE sequences from another time point or anatomical compartment, (ii) 2 matches were observed between STIP-Seq blood T0 sequences and HIV-PULSE lymph nodes T1 sequences (1 intact and 1 PSI/MSD-defective), but did not overlap with HIV-PULSE sequences from another time point or anatomical compartment (Supplementary Fig. 6). Importantly, one blood T0 STIP-Seq sequence with a large deletion that does not affect p24 production overlapped with HIV-PULSE sequences retrieved across all time points and anatomical compartments (MRC21, Supplementary Fig. 6). The remaining matches between STIP-Seq and HIV-PULSE sequences all comprised PSI/MSD-defective proviruses and were shared between blood T0 and T1 ($n = 3$), blood T0 / lymph nodes T1 ($n = 2$), and blood T0 / blood T1 / lymph nodes T1 ($n = 2$). In conclusion, our results indicate that clonally infected cells carrying inducible proviruses are capable of persisting over time and can disseminate across various anatomical compartments.

Discussion

To our knowledge, this is the first study that simultaneously analyzes the phenotypic traits of the translation-competent reservoir together with the near full-length proviral landscape in matched blood, lymph node and gut samples from ART-suppressed individuals. While the phenotype of p24+ cells in matched blood and lymph nodes has been studied in the context of viremic infection[41], assessing it in ART-suppressed individuals has been challenging due to the lack of potent and specific LRAs to boost viral production in latently infected cells. In this study, we took advantage of the previously reported Tat-LNP/PNB combination to maximize latency reversal[18], thereby enhancing our capacity to detect, quantify and phenotypically characterize translation-competent reservoir cells in matched blood and lymph nodes. In agreement with a previous study that reported similar frequencies of viral RNA-expressing CD4 T cells in blood and lymph nodes following latency reversal[23], we observed that frequencies of p24-expressing cells along with intact/total HIV copies ratios were comparable between the two anatomical sites.

The CD4 T cell memory compartment has traditionally been subdivided into TCM and TEM cells[42]. TCM recirculate between lymphoid organs and the blood, and have the ability to quickly proliferate upon antigen re-encounter. In contrast, TEM cells primarily circulate through the blood and target inflamed tissues, where they quickly release effector molecules to provide immediate defense. Accordingly, we showed that p24+ cells from the blood were found in both TCM/TTM and TEM subsets, while the vast majority of p24+ cells from the lymph nodes displayed a TCM/TTM phenotype and were barely observed in the TEM fraction.

In lymphoid tissues, Tfh cells are characterized by the expression of CXCR5, which directs them to B-cell follicles, as well as by PD1 and ICOS. The minor fraction of Tfh cells that resides in germinal centers is defined by high expression levels of CXCR5 and PD1[43]. While Tfh cells are classically found in secondary lymphoid organs, a small population of similar cells is also present in the blood, commonly termed circulating Tfh[44,45]. In our study, most blood-derived p24+ cells did not express CXCR5, whereas around 50% of lymph node-derived p24+ cells expressed this molecule, suggestive of their residency or capacity to migrate to B cell follicles. This observation is in line with recently published data showing that HIV-infected cells in lymph nodes from individuals treated for 10-15 years are enriched in CXCR5-expressing cells[13]. However, our study indicates that CXCR5-expressing p24+ cells do not co-express high levels of PD1, suggesting that GC Tfh cells do not harbor the translation-competent reservoir in lymph nodes from long-term treated individuals. Consistent with this observation, several findings suggest that the contribution of GC Tfh cells to the HIV-1 reservoir may diminish with the duration of viral suppression: (i) the frequency of GC Tfh cells progressively declines with time on ART[29], potentially influencing their contribution to the viral reservoir. In agreement with this observation, participants treated for more than 10 years in our study displayed low frequencies of GC Tfh cells; (ii) levels of cell-associated HIV RNA detected in the PD1-positive subset from lymph nodes inversely correlate with the treatment duration[29], suggesting a progressive elimination of the viral reservoir in this population.

Cytotoxic CD4 T cells are primarily identified within highly differentiated cell subsets such as TEM and TTd cells[46] and have been shown to expand in response to chronic viral infections such as CMV, HIV-1 and hepatitis[47]. In our study, frequencies of GZMA+ CD4 T cells were highly variable between participants in the blood, comprising up to 45% of total CD4 T cells in one of the study participants. As previously shown by our team in the context of Tat-LNP stimulation alone[18], GZMA was preferentially expressed by peripheral p24+ cells following Tat-LNP/PNB treatment. In contrast, translation-competent reservoir cells from the lymph nodes did not express GZMA, suggesting that GZMA expression does not represent a distinctive feature of the inducible reservoir in lymph nodes. Rather, the lack of GZMA expression by p24+ cells seems to reflect the limited expression of GZMA observed in CD4 T cells from the lymph nodes, consistent with the observation that cytolytic T cells are retained within the intravascular circulation[48].

Finally, Bcl2 was detected in nearly all p24+ cells in the blood and lymph nodes, consistent with the majority of CD4 T cells being Bcl2+ in these anatomical sites. As opposed to previously reported data[32], blood-derived p24+ cells did not exhibit a higher mean fluorescence intensity for Bcl2 compared to p24- cells, which could potentially be explained by the choice of LRAs used to reactivate HIV from latency (PMA/ionomycin vs Tat-LNP/PNB).

Our observations support the notion that preferential expression of certain markers by p24+ cells mirrors the general phenotype of CD4 T present in distinct anatomical locations, rather than indicating unique intrinsic features that would support their long-term persistence regardless of their anatomical site. Antigen-specificity may also influence the phenotype and the anatomical location of p24+ cells. For instance, previous studies have described intrinsic differences between CMV- and HIV-specific T cells. CMV-specific T cells typically exhibit a TEM cytolytic profile[34,48,49] and are primarily found in the blood[50,51], whereas HIV-specific T cells have reduced cytotoxic responses relative to CMV-specific T cells[34,48] and are found more broadly across anatomical compartments[48,51,52]. Future studies analyzing the TCR of p24+ cells may help us determine how antigen stimulation shapes the phenotype and migration properties of p24+ cells.

Immunophenotyping documentation using the cell sorter while simultaneously sorting $CD3^{+}CD4^{+/-}$ T cells enabled to obtain both immunological and virological data from the same cell cultures despite the low cell input. We showed that HIV-PULSE can be successfully used to characterize the proviral composition of the HIV-1 reservoir not only in blood but also in tissues, even on paraformaldehyde-fixed samples. This high throughput approach allowed us to recover a total of 3421 near full-length proviral genomes across blood T0/T1, lymph nodes T1 and gut T1. Proviral genome classification revealed that similar fractions of HIV-1 sequences are putatively intact between the two longitudinal time points in the blood and the lymph nodes. The lower fraction of intact sequences detected in the gut might be related to the low number of immune cells obtained in this compartment starting from 15 micro biopsies. Interestingly, in MRC21, we identified a mix of CCR5 and CXCR4-tropic viruses, which may offer an interesting explanation for the high fraction of p24+ cells with a naïve T cell phenotype in this participant.

We observed several clones of infected cells at each time point and in all anatomical compartments, confirming the overall clonal nature of the HIV-1 reservoir in ART-suppressed individuals. The variation in clonality estimates among participants indicates that the contribution of infected cell clones to the viral population can differ on an individual level but remains nonetheless a common feature of the HIV reservoir. We also found that the estimate of clonality was systematically lower in the lymph nodes compared to the blood for each participant included in this study. As previous research has demonstrated that clonal expansion is a more prevalent phenomenon in the effector memory fraction than in less differentiated phenotypes[53,54], the lower degree of clonality in the lymph nodes may be partly explained by the predominance of central memory cells among the infected cells in this anatomical compartment. In the blood, multiple infected clones overlapped between the two longitudinal time points, highlighting their long-term persistence. Of note, in MRC15, 95% of the p24+ cells in the blood at T0 carried a virus integrated into *SNTB1* (Supplementary Data 2). While this specific clone was detected in multiple HIV-PULSE replicates within the blood at T0, it was never detected in the blood nor in the tissues at T1. Therefore, and as previously reported[55,56], while some clones of infected cells persist over many years on ART, others tend to wane, including those carrying an inducible provirus.

Out of the 319 infected clones detected, we observed 115 overlapping clones between the blood and tissues, including three clones with an intact provirus, suggesting that clones disseminate in the body through the blood. Interestingly, out of 55 infected clones detected in the gut, 37 were shared with the blood. Several factors may explain why some sequences found in the gut were not detected in the blood: (i) the detection of four infected clones that were exclusively shared between gut and lymph nodes without being detected in the blood indicates that sampling bias may play a role; (ii) the direct infection of gut CD4 T cells with a resident memory phenotype, which comprise around 2-10% of the total CD4 T cells in this anatomical site[57], may further participate to the lack of proviral overlap between blood and tissues. The simultaneous sampling of blood and tissues in our study complements two previous post-mortem studies where NFL sequencing analysis was performed in tissue samples but no blood sampling before death was conducted[6,7]. Interestingly, while certain clones of infected cells overlapped between blood and lymph nodes, p24+ cells from these two compartments appeared to have distinct phenotypes. This observation suggests that cells with an identical provirus will display different phenotypes depending on their anatomical location, confirming the previously reported plasticity of immune cells in response to their microenvironment[58,59].

Finally, APD calculations, panmixia testing and phylogenetic analyzes all indicated a lack of compartmentalization between the various time points and anatomical sites. The absence of viral diversification between blood and matched tissues seems to indicate that viral replication does not occur in the tissues[20,26,27,39], despite previous reports of low ARV penetration in certain anatomical sites[1,2]. Although this observation has been documented previously[20,26,27,39], our study stands out by the extensive number of full-length proviral sequences obtained, further reinforcing the validity of previous findings. Additionally, we identified a few links between STIP-Seq sequences retrieved from blood at T0 and HIV-PULSE sequences from blood and tissues at T1, indicating that certain clones of infected cells carrying an inducible provirus persist over time and may participate to viral rebound following ART cessation.

Our study has some limitations. The limited number of cells retrieved from tissues and their apparent fragility during culture and stimulation conditions restricted the quantity of p24+ cells and HIV-1 proviral sequences we could obtain from lymph node and gut samples compared to the blood compartment. Simulation data have shown that 20 sequences per anatomical site is an adequate sample size to make confident claims about compartmentalization[60]. In our study, more than 20 distinct HIV-1 sequences were obtained in 9/9 participants in the blood, 3/5 participants in the lymph nodes, and in 4/8 participants in the gut. Primer design constraints and amplification biases towards shorter fragments in long-range PCR amplification, as previously discussed by our team[35] and others[61,62], may have hindered the detection of additional overlaps. Furthermore, while the expression of most markers included in this analysis was preserved following Tat-LNP/PNB stimulation, the frequency of some specific subsets such as GZMA+ cells and GC Tfh cells was somewhat reduced in one or more participants. Thus, additional studies are required to confirm that GC Tfh cells are not a preferential HIV reservoir in long-term treated individuals. Additionally, the upregulation of T cell residency markers such as CD69 in response to panobinostat precluded the analysis of resident memory CD4 T cells in our study. While parallel phenotypic and virological characterization was successfully performed and suggests that clones of infected cells display distinct phenotypes based on their anatomical location, our study does not provide direct proof. To answer this question definitively, single-cell studies that can identify clones of infected cells shared between anatomical sites while collecting information on the phenotype of these cells are needed.

Overall, using a combination of innovative and high throughput approaches, our study provides additional insights into the phenotypic profile of the translation-competent reservoir and the reservoir's viral composition in blood and matched tissues from ART-suppressed individuals. The observation that the inducible reservoir exhibits a different phenotype between blood and lymph nodes suggests that different interventions may be needed to eliminate the reservoir in these two compartments. Furthermore, our findings indicate that GC Tfh cells may not be preferential targets for cure interventions in long-term treated individuals. Lastly, our study reaffirms that ongoing replication does not appear to occur in tissues under ART, and that HIV-infected cells disseminate across various organs using the blood as a conduit.

Methods

Ethics statement

All participants were adults and provided written informed consent. The Mercuri study was approved by the Ethics Committee of the Ghent University Hospital (Belgium) (EC number: Mercuri BC-07056, NCT04305665[18]). Analyzes presented in this manuscript are covered by the study protocol. There was no experimental drug administration performed in this study. Participants were compensated 230 euros for the extensive sampling (blood draw, leukapheresis, colonoscopy and lymph node resection).

Study population

A total of $n = 9$ people living with HIV on stably suppressive ART were included in this study (Supplementary Table 1). Participants were recruited at Ghent University Hospital. 1/9 individuals is female, 8/9 are male; the limited representation of female individuals in our study is a direct reflection of the population of HIV-1 seropositive individuals in Belgium, which predominantly consists of men who have sex with men. Gender was determined based on self-report.

Antibodies

Fixable Viability Stain 510 was obtained from ThermoFisher Scientific (#L34957, 1/1000). The following antibodies were used in staining experiments: CD3 AF700 Clone UCHT1 (BD Biosciences, #557943, 1/50), CD4 BV786 Clone SK3 (BD Biosciences, #563881, 1/200), CD8 BV510 Clone RPA-T8 (BioLegend, #301047, 1/200), CD45RO PE Clone HI100 (BD Biosciences, #5555493, 1/25), CD27 BV605 Clone L128 (BD Biosciences, #562656, 1/100), PD1 BB700 Clone EH12.1 (BD Biosciences, #566461, 1/100), CXCR5 BV421 Clone J252D4 (BioLegend, #356919, 1.5/100), CD45 FITC Clone HI30 (BioLegend, #304038, 1/50),

ICOS BV711 Clone C398.4 A (BioLegend, #313547, 1/50), GZMA PE-Cy7 Clone CB9 (BioLegend, #507221, 1/50), Bcl2 PE-CF594 Clone Bcl-2/100 (BD Biosciences, #563601, 1/100). For p24 staining, we used a combination of two antibodies: p24 KC57-FITC (Beckman Colter, #6604665, 1/500) and p24 28B7-APC (MediMabs, #MM-0289-APC, 1/400).

Cell isolation from blood draws

8 tubes of 9 mL blood were collected per participant and processed within two hours after collection. Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll density gradient centrifugation, before proceeding to CD4 isolation and freezing of the cells in FCS/DMSO 10%.

Cell isolation from lymph nodes

The inguinal lymph node was harvested and placed in a 50 mL tube containing complete RPMI (supplemented with 10% FCS and 1% Pen/Strep), then kept on ice until processing, which occurred within two hours of collection. After manually removing the surrounding fat, the lymph node was cut into small pieces. Each piece was mechanically disrupted on a 70 μm cell strainer placed over a 50 mL tube using tweezers. The strainer was rinsed multiple times with complete RPMI, and the lymph node mononuclear cells (LNMCs) were centrifuged for 10 minutes at 1700 rpm. The resulting pellet was resuspended in PBS, and any remaining fat was eliminated by passing the cells through a 40 μm cell strainer. Cell counting was performed using Turk's solution before proceeding to CD4 isolation and freezing of the cells in FCS/DMSO 10%.

Cell Isolation from gut microbiopsies

40 microbiopsies (10 from rectum, 15 from colon, 15 from ileum) per participant were harvested and placed in 3×15 mL tubes containing 10 mL complete RPMI (supplemented with NEAA 1% (Gibco, #11140050) and Sodium pyruvate 1% (Gibco, #11360070)), then kept on ice until freezing, which occurred within two hours of collection. Microbiopsies were transferred in cryovials (max 5 per cryovial) containing 1 mL of Cryostor CS10 (StemCell Technologies, #100-1061). After thawing in a 37 °C warm water bath, 15 microbiopsies (5X rectum, 5X colon, 5X ileum) were gathered on a 100 μm cell strainer and sequentially rinsed with complete RPMI and PBS/FBS 2%. The partial elimination of the epithelial cell layer was achieved through mechanical disruption in 5 mL HBSS 1X supplemented with 5 mM EDTA (15 min, 37 °C). To facilitate tissue digestion and achieve a single-cell suspension, microbiopsies underwent additional enzymatic digestion in 5 mL RPMI 25 mM HEPES supplemented with Collagenase IV and VIII (Sigma, #C51381 and #C2139, respectively) (1 mg/mL each, 30 min, 37 °C). The single-cell suspension was passed through a 40 μm cell strainer and enzymatic digestion was blocked by the addition of cold PBS. Following cell centrifugation, enrichment of $CD45^+$ immune cells was achieved using the EasySep™ Human EpCAM Positive Selection Kit II (#17846). Following enrichment, CD45-enriched cells were resuspended in 1 mL PBS/FBS 2%: 10 μL was employed to assess the purity by flow cytometry (Supplementary Fig. 1), 50-100 μL were used for the complete staining procedure, and the remaining volume was utilized to make a dry pellet for downstream virological assays. Of note, the presence of debris and residual epithelial cells in the CD45-enriched fraction from the gut prevented accurate cell counting.

Isolation of CD4 T cells

CD4 T cells were isolated from PBMCs and LNMCs by negative magnetic selection using the EasySep Human CD4 T Cell Enrichment Kit (StemCell Technology, #19052). Purity was typically > 98% following CD4 isolation from PBMCs. Frequencies of CD3-negative T cells typically ranged between 2.3% and 26.4% (median = 4.9%) following CD4 enrichment from LNMCs.

HIV-flow procedure

Between 1.7×10^6 and 5×10^6 CD4 T cells from blood and lymph nodes were resuspended at 2×10^6 cells/mL in complete RPMI supplemented with antiretroviral drugs (200 nM raltegravir, 200 nM lamivudine) to avoid new cycles of replication. Cells were rested for at least an hour at 37 °C before being stimulated with 50 nM panobinostat (Selleckchem, #LBH589) and Tat-LNP (250 ng/mL; 1.4 nM). A non-stimulated condition was run in parallel for each sample in order to check for the impact of Tat-LNP/PNB stimulation on the phenotype of the cells. Following LRA stimulation, frequencies of p24-producing cells were measured by using a combination of 2 antibodies targeting the p24 protein (p24 KC57-FITC, p24 28B7-APC), as previously described[11]. The detailed protocol of the HIV-Flow procedure can be found here: https://doi.org/10.17504/protocols.io.w4efgte. Cells were acquired on the BD FACSAria Fusion Cell Sorter. While recording, $CD3^+/CD8^-/CD4^{+/-}$ T cells were sorted into a 15-mL tube and then centrifuged at high speed to make a dry pellet. This procedure enabled the phenotypic characterization and quantification of the translation-competent reservoirs in the blood and lymph nodes, while recovering CD4 T cells for downstream virological assays. The total number of CD4 T cells analyzed by HIV-Flow per participant is displayed in Supplementary Table 3.

Cell lysis

Pellets of sorted $CD3^+CD4^{+/-}$ T cells from blood and lymph nodes (between 0.4 and 1.5 × 10^6 cells/pellet), as well as pellets of CD45-enriched cells from the gut were lysed in 100 μL directPCR lysis reagent (Viagen Biotech; #301-C). Cells were incubated for 16 h at 55 °C, followed by 45 min at 85 °C. Lysed cells were stored at −20 °C until they were used.

Intact proviral DNA assay

The Rainbow proviral HIV-1 DNA dPCR assay was performed on a QIAcuity Four digital PCR platform (Qiagen, Germany), as previously described[37]. The total number of CD4 T cells assayed by IPDA per participant is displayed in Supplementary Table 3. 8 μL of lysed CD4 T cells from blood and lymph nodes (32,000-120,000; mean = 72,200 cells per reaction) and 5 μL of lysed CD45-enriched cells from gut (cells were not counted) were used as input per replicate. Triplicates were performed. The reaction mix preparation for the Rainbow proviral HIV-1 DNA assay was as follows: 10 μL 4x concentrated Qiacuity Probe Master Mix (Qiagen, #250102,), 2 μL of each primer/probe set (final concentrations for RU5: 0.675 μM primers/0.187 μM probe, PSI: 0.675 μM primers/0.187 μM probe, Env: 0.5 μM primers/0.25 μM probe; Supplementary Table 4), 0.3 μL restriction enzyme XBaI (100,000 units/ml), 5-8 μL lysed cells and nuclease free water to 40 μL total volume per reaction. The PCR reaction mixes were transferred into a 26k 24-well nanoplate (Qiagen, #250001). Partitioning of the PCR mix was performed on QIAcuity Four digital PCR system after sealing the plate. The digital PCR program started with 2 min at 95 °C for the activation of the Qiacuity probe mix, followed by 40 cycles of 94 °C for 30 sec and 56 °C for 60 sec. Imaging settings were 500 ms, 500 ms, 400 ms, 200 ms and 400 ms in the green, yellow, orange, red and crimson channel, respectively.

HIV-PULSE assay

A modified version of the published HIV-PULSE assay was performed on blood and tissue samples collected at T1. The protocol, including changes to enzymes used for pre-amplification, sequencing kits and software, is described below, and can be found here: https://doi.org/10.17504/protocols.io.8epv5rby4g1b/v1. Changes to the protocol improved the UMI bin centroid accuracy for bins with a coverage of at least 15 reads from 99.91% (R10.3, LSK109, Guppy v.5.0.17, pre-amplification with LongAmp Taq DNA Polymerase) to 99.97% (R10.4.1, LSK114, Guppy v.6.5.7, pre-amplification with PrimeSTAR GXL DNA

Polymerase) (Supplementary Fig. 7A). The list of primers is included in Supplementary Table 4. The total number of CD4 T cells assayed by HIV-PULSE per participant is displayed in Supplementary Table 3.

Pre-amplification. A first PCR using a high fidelity enzyme was used to specifically target and pre-amplify HIV-1 proviral templates using the outer primers of a nested HIV-1 primer set. Each PCR reaction contained lysed genomic DNA, 1 µl of PrimeSTAR GXL DNA Polymerase (Takara Bio, #R050B), 0.2 µM of each primer (First PCR F, First PCR R), 4 µl of dNTP mixture, 10 µl of 5 × 5X PrimeSTAR GXL Buffer in a final volume of 50 µl. The following cycling conditions were used: 98 °C for 2 min; 6 cycles (10 sec at 98 °C, 15 sec at 65 °C, 10 min at 68 °C); 68 °C for 10 min. PCR products were cleaned using CleanPCR magnetic beads (CleanNA, #CPCR-0050) at a 1.0× beads:sample ratio.

Tagging HIV-1 templates. A second PCR was performed to tag both ends of the pre-amplified proviral HIV-1 templates with a tailed UMI. Primers were designed to contain: (i) a synthetic primer binding site used in later stages for amplification, (ii) a UMI with a repetitive pattern of 12 random nucleotides and 6 degenerate nucleotides (Y/R) and (iii) an HIV-1 inner primer of the nested primer set to target the pre-amplified templates. Each PCR reaction contained all the cleaned pre-amplified product (30 µl), 2 µl of LongAmp Taq DNA Polymerase (NEB, #M0323L), 0.5 µM of each primer (Second PCR F UMI, Second PCR R UMI), 1.5 µl of 10 mM dNTPs (Promega, #C1141), 10 µl of 5× LongAmp Taq Reaction Buffer in 50 µl. The following cycling conditions were used: 94 °C for 1 min 15 sec; 2 cycles (94 °C for 30 sec, 58 °C for 30 sec, 65 °C for 10 min); 65 °C for 10 min. Tagged PCR products were cleaned using CleanPCR magnetic beads (CleanNA, #CPCR-0050) in a custom buffer solution (based on the 'SPRI size selection protocol for >1.5–2 kb DNA fragments' protocol provided by Oxford Nanopore Technologies (ONT)) at a 0.9× beads:sample ratio and eluted in 30 µl of nuclease-free water.

Amplification of UMI-tagged proviruses. The next steps used four consecutive PCR amplification rounds each of 10 cycles followed by a cleanup to produce enough template input required for long-read sequencing while preserving amplicon size distributions. Here, we made use of a primer set that binds to the synthetic binding site incorporated during the previous tagging stage. The PCR mix consists of 2 µl of LongAmp Taq DNA Polymerase (NEB, #M0323L), 0.5 µM of each primer (ncec_pcr_fw_v7, ncec_pcr_rv_v7), 1.5 µl of 10 mM dNTPs (Promega, #C1141), 10 µl of 5× LongAmp Taq Reaction Buffer in 50 µl. For the first PCR amplification round all the cleaned tagging products from the previous step (30 µl) were used as template inputs, whereas only 10 µl of the cleaned product of the previous round were used during the second, third and fourth amplification rounds. The following cycling conditions were used: 94 °C for 1 min 15 sec; 10 cycles (94 °C for 30 sec, 58 °C for 30 sec, 65 °C for 10 min); 65 °C for 10 min. PCR products were cleaned after each consecutive round using regular CleanPCR magnetic beads (CleanNA, #CPCR-0050) at a 1.0× beads:sample ratio and eluted in 30 µl of nuclease-free water. During the last round of 10 cycles, the regular primers were switched for a custom set of tailed primers to barcode the PCR products from the same participant with a specific, identical identifier (Supplementary Table 4). After the last PCR round, the end products were visualized on 1% agarose gel and the DNA concentration was determined using a Qubit 3.0 fluorometer with the Qubit dsDNA BR assay kit (ThermoFisher Scientific, #Q32853).

ONT long-read sequencing. Samples were multiplexed using the Native Barcoding Kit 24 V14 kit (ONT, #SQK-NBD114-24) using the following strategy: each PCR replicate was assigned to a different ONT barcode and contained equimolarly pooled PCR products from different participants. In later stages, this allowed the assignment of reads to the correct PCR replicate by the ligated ONT barcode and to the correct sample by the participant-specific identifier attached during the last PCR round. For library preparation, the Native Barcoding Kit 24 V14 kit (ONT, #SQK-NBD114-24) was used following the manufacturer's instructions. Samples were sequenced on a MinION ONT device using MinION R10.4.1 flow cells (ONT, #FLO-MIN114) and the MinKNOW v.23.04.5 software followed by basecalling at super accuracy mode and demultiplexing with Guppy v.6.5.7.

Bioinformatics analysis of long-read data. For the analysis of long-read data, the HIV-PULSE UMI data analysis was used (https://github.com/laulambr/longread_umi_hiv)[35]. The main changes include updated software versions of samtools (v.1.11), medaka (v.1.9.1) and racon[63] (v.1.4.20). Before analyzing the data using the UMI pipeline, the demultiplexed ONT reads were mapped against the HXB2 reference sequence using minimap2 (2.17) to filter out non-HIV-1 reads. Next, the 'longread_umi nanopore_pipeline' workflow was run separately on each replicate read dataset using the following settings: -s 200 -e 200 -m 1500 -M 10000 -f CAAGCAGAAGACGGCATACGAGAT -F AAGTAGTGTGTGCCCGTCTGTTGTGTGAC -r AATGATACGGCGACCACCGAGATC -R GGAAAGTCCCCAGCGGAAAGTCCCTTGTAG -c 3 -p 1 -q r103_hac_g507 -U 'r103_min_high_g360'. The workflow consists of the following consecutive steps: (i) trimming and filtering of the HIV-1 long-read sequencing data using Porechop (v.0.2.4, https://github.com/rrwick/Porechop), Filtlong (v.0.2.0, https://github.com/rrwick/Filtlong) and cutadapt (v.2.7); (ii) extraction of UMI reference sequences using cutadapt (v.2.7) and usearch[64]; (iii) binning of reads to UMI combinations using bwa (v.0.7.17) and samtools (v.1.11) while excluding chimeric artifacts; (iv) generation of bin centroid sequences using usearch and minimap2 (v.2.17) and (v) polishing of bin centroid data by multiple rounds of racon[63] (v.1.4.20) and a final round of Medaka (v.1.9.1, https://github.com/nanoporetech/medaka). Next, a custom bioinformatics workflow specific to the HIV-PULSE protocol was run as described before[35] to correct for pre-amplification, improve final bin accuracy and evaluate clonality among PCR replicates. Supplementary Fig. 7B shows the HIV-PULSE genome accuracy for the samples included in this study. Among 37 genomes that were overlapping between HIV-PULSE and STIP-Seq, 29 HIV-PULSE sequences had 100% agreement with their STIP-Seq reference counterpart. For the remaining 8 sequences, minor discrepancies (an average of 1.6 base pair substitutions) were observed. These discrepancies did not affect the proviral classification or clonal linkage.

STIP-Seq assay

CD4 T cell stimulation. CD4 T cells from blood T0 were resuspended at 2×10^6 cells/mL in RPMI + 10% Fetal Bovine Serum and antiretroviral drugs were added to the culture (200 nM raltegravir, 200 nM lamivudine). Cells were rested for at least an hour at 37 °C before being stimulated with the following LRAs: (i) 1 µg/mL ionomycin (Sigma, #I9657) and 162 nM PMA (Sigma, #P8139), (ii) 50 nM panobinostat (Selleckchem, #LBH589) and Tat-LNP (250 ng/mL; 1.4 nM), (iii) 162 nM PMA (Sigma, #P8139) and Tat-LNP (250 ng/mL; 1.4 nM) (Supplementary Data 2).

Methanol-based HIV-Flow procedure for STIP-Seq. The methanol-based HIV-Flow procedure was performed as previously described by Pardons et al. [18], without further modification. The detailed protocol can be found here: https://protocols.io/view/methanol-based-hiv-flow-bpedmja6.

Multiple displacement amplification (MDA). Whole genome amplification of single sorted cells was carried out by multiple displacement amplification with the REPLI-g single cell kit (Qiagen, #150345), according to manufacturer's instructions. A positive control, consisting of 10 p24- cells sorted into the same well, was included on every plate.

Quantitative polymerase chain reaction (qPCR) for *RPP30*. After whole genome amplification by MDA, reactions were screened by qPCR on the *RPP30* reference gene (primers described in Supplementary Table 4), as described previously[36]. Reactions that yielded a cycle of threshold (Ct) value of 38 or lower, were selected for further downstream processing.

Integration site analysis. MDA reactions that were positive for *RPP30* were subjected to integration site sequencing by a modified version of the integration site loop amplification (ISLA) assay (primers described in Supplementary Table 4), as described previously[36]. Resulting amplicons were visualized on a 1% agarose gel and positives were sequenced by Sanger sequencing. Analysis of the sequences was performed using the 'Integration Sites' webtool (https://indra.mullins.microbiol.washington.edu/integrationsites).

Near full-length proviral genome amplification. MDA wells that were *RPP30* positive were subjected to 5' and 3' half genome amplification. The 25 μL PCR mix for the first round is composed of: 5 μL 5X Prime STAR GXL buffer, 0.5 μL PrimeStar GXL polymerase (Takara Bio, #R050B), 0.125 μL ThermaStop (Sigma Aldrich, #TSTOP-500), 250 nM forward and reverse primers, and 1 μL of 1/5 diluted MDA product. The mix for the second round has the same composition and takes 1 μL of the first-round product as an input. Thermocycling conditions for first and second PCR rounds are as follows: 2 min at 98 °C; 35 cycles (10 sec at 98 °C, 15 sec at 62 °C, 5 min at 68 °C); 7 min at 68 °C. The primer sequences are summarized in Supplementary Table 4. The summary of all PCRs that were performed is shown in Supplementary Data 2. Near full-length proviral genome sequencing and assembly were performed as previously described[36].

Proviral genome classification

NFL proviral genome classification for STIP-Seq and HIV-PULSE sequences was performed using the Proviral Sequence Annotation & Intactness Test (ProSeq-IT) (https://psd.cancer.gov/tools/tool_index.php), HIVIntact[65] and the publicly available 'Gene Cutter' and 'Hypermut' webtools from the Los Alamos National Laboratory HIV sequence database (https://www.hiv.lanl.gov). Proviral genomes were classified in the following sequential order: (i) 'Inversion', the presence of internal sequence inversion, defined as region of reverse complementarity. (ii) 'Large internal deletion', internal sequence deletion of >1000 bp. (iii) 'Hypermutated', APOBEC-3G/3F-induced hypermutation. For each participant, a consensus sequence was generated from the remaining genomes and a new alignment with the consensus sequence as the reference was created. This alignment was then analyzed using the 'Hypermut' webtool, with proviruses scoring $P < 0.05$ considered hypermutated. (iv) 'Packaging signal and/or major splice donor (PSI/MSD) defect', proviruses containing a deletion > 7 bp found in any of the four stem loops of the PSI region (SL1 (HXB2: 691–734), SL2 (HXB2: 736–754), SL3 (HXB2: 766–779) and SL4 (HXB2: 790–810)). This also includes the absence and/or point-mutation of both the MSD site (sequence GT, HXB2: 744–745) and the cryptic donor site (sequence GT, HXB2: 748–749). Proviruses with a deletion covering PSI/MSD that extended into the *gag* gene, thereby removing the *gag* AUG start codon, were also classified into this category. (v) 'Premature stop-codon/frameshift', premature stop-codon or frameshift caused by mutation and/or sequence insertion/deletion in the essential genes *gag*, *pol* or *env*. Proviruses with insertion/deletion >49 nt in *gag*, insertion/deletion >49 nt in *pol*, or insertion/deletion >99 nt in *env* were also classified into this category. (vi) 'Intact', proviruses that displayed none of the above defects were classified into this category.

Testing of tropism, diversity and panmixia

Proviral sequences were aligned against the HXB2 reference sequence using MAFFT (v.7.471)[66]. For further analysis, only distinct sequences were used for each dataset (blood T0/T1, lymph nodes T1 and gut T1), collapsing clonal sequences into a single representative sequence per dataset. For tropism prediction, only proviral sequences with full coverage of the V3 region in *env* were retained (while excluding proviruses classified with inversions or hypermutations). Prediction was performed via WebPSSM[67] and Geno2pheno[68] coreceptor v2.5. For diversity analysis, only proviral sequences with full coverage of the *nef* gene were retained (while excluding proviruses classified with inversions or hypermutations). DIVEIN[69] was used to calculate average pairwise distance among proviral sequences for each different dataset. The hyphy software package (v.2.5.62)[70] was used to conduct tests for panmixia based on the *nef* alignments (Supplementary table 2). A Hudson, Boos, and Kaplan's nonparametric test for population structure was employed as a distance-based method. The K_{ST} values range from 0 (no compartmentalization) to 1 (full compartmentalization) and are tested for statistical significance via randomization test using 1000 permutations. The Slatkin-Maddison (SM) test was used a tree-based method to assess panmixia. First, a phylogenetic tree of each *nef* alignment was created with IQ-TREE v.1.6.11[71], using the built-in model finder and ultrafast bootstrapping with 1000 replicates. Next, a standard SM test was run between each dataset using 1000 replicates providing the minimum number of migration events needed to explain the observed distribution of traits on a phylogenetic tree. The *p*-value indicates whether the observed distribution is significantly different from what would be expected under a random distribution. For a dataset to be considered compartmentalized, results from both the distance- and tree-based methods needed to have a significant *p*-value.

Clonal classification

HIV-PULSE. Clonal classification using the HIV-PULSE assay was performed following the methodology outlined in the original publication[35]. To link sequences across timepoints and compartments, >99.99% sequence identity following MAFFT alignment[66] was required, ensuring robust identification of clones.

STIP-Seq. p24+ cells were defined as clonal if proviruses shared 100% sequence identity, with further validation provided by matched integration sites.

Integrating HIV-PULSE and STIP-Seq data. For each participant, proviral genomes from STIP-Seq and HIV-PULSE were aligned to HXB2 using MAFFT (v.7.471)[66]. These alignments were manually inspected, and clones were identified based on the following criteria:

1. We considered a proviral overlap between HIV-PULSE and STIP-Seq proviruses if sequences were 100% identical.
2. If an exact match was not found, we still considered a proviral overlap between HIV-PULSE and STIP-Seq proviruses if sequence identity was greater than 99.5% and the same deletion junction sites were observed.

Data representation and statistical analyzes

Dot plots and pi charts were generated with Graphpad Prism (v.8.0.2). Flow cytometry analyzes involving manual gating were done with FlowJo v.10.6.2. FlowSOM analyzes were performed using FlowJo v.10.10.0, R v.4.3.2, and the following plugins: FlowSOM (v.3.0.18), cluster explorer (v.1.7.6) and UMAP (v.4.0.4). All plugins were downloaded on FlowJo Exchange. FlowSOM analysis on non-stimulated cells (Supplementary Fig. 2C) included 1600 $CD4^+$ T cells per participant and per anatomical compartment (blood, lymph nodes and gut; T1), for a total of 30,400 cells (9600, 8000, 12,800 cells from blood, lymph nodes and gut, respectively). FlowSOM analysis on Tat-LNP/PNB-stimulated cells (Fig. 2F) included 5000 $CD3^+CD4^{+/-}$ T cells per participant and per anatomical compartment (blood and lymph nodes; T1), along with 128 blood-derived p24+ cells and 55 lymph node-derived p24+ cells, for a total of 50,183 cells.

Proviral sequences from HIV-PULSE and STIP-Seq were aligned against the HXB2 reference sequence using MAFFT v.7.471[66]. Phylogenetic trees were constructed using IQ-TREE v.1.6.11[71] using the built-in model finder and ultrafast bootstrapping with 1000 replicates. R v.4.3.2, ggplot v.2.3.5 and ggtree v.3.10.1 were used for visualization and annotation of trees.

No statistics were run for Fig. 2A–C and Supplementary Fig. 3A-B, as the minimum sample size for the Wilcoxon signed-rank test to detect a statistically significant difference (at $\alpha = 0.05$) is considered 6 paired datapoints. For Figs. 3C, E and 4B: the following groups were compared two by two: B_T0 vs B_T1, B_T1 vs L_T1, B_T1 vs G_T1, L_T1 vs G_T1. For pairwise comparisons between groups with normal residuals (Figs. 3E and 4B): two-sided parametric paired t-tests were performed (missing values were discarded). Additionally, a mixed-effects analysis was also performed, as this model can handle missing values. For pairwise comparisons between groups with non-normal residuals (Fig. 3C), two-sided non-parametric Wilcoxon tests were performed (missing values were discarded). All statistical tests and *p*-values obtained are displayed in Supplementary Data 1.

Data availability

The HIV-1 near full-length sequences generated in this study by HIV-PULSE and STIP-Seq have been deposited in the GenBank database under accession codes PV026554-PV028940. The HIV-PULSE and STIP-Seq sequences generated in previous papers and used in this study can be found in the GenBank database under accession codes OR245577–OR246884 and OQ596882-OQ596960. Tables exceeding an A4 page size are provided as Supplementary data files (Supplementary Data 1, 2). Source data are provided with this paper and on figshare (https://doi.org/10.6084/m9.figshare.26780080). For all requests pertaining to the utilization of the Tat-LNP molecule, kindly direct your inquiries to Janssen's Infectious Diseases and Diagnostics Department (Beerse, Belgium). Contact person: Johan Vingerhoets. Source data are provided with this paper.

Code availability

Scripts used in this study are freely available on GitHub: https://github.com/laulambr/longread_umi_hiv[72], https://github.com/laulambr/virus_assembly[73]. A description of the key operations and instructions on how to install and run the code are provided on the GitHub page. In addition, a test dataset and a description of expected outputs with expected run times are provided.

References

1. Fletcher, C. V. et al. Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations in lymphatic tissues. *Proc. Natl Acad. Sci. USA* **111**, 2307–2312 (2014).
2. Fletcher, C. V. et al. Persistent HIV transcription and variable antiretroviral drug penetration in lymph nodes during plasma viral suppression. *AIDS (Lond., Engl.)* **36**, 985–990 (2022).
3. Connick, E. et al. CTL fail to accumulate at sites of HIV-1 replication in lymphoid tissue. *J. Immunol. (Baltim., Md.: 1950)* **178**, 6975–6983 (2007).
4. Estes, J. D. et al. Defining total-body AIDS-virus burden with implications for curative strategies. *Nat. Med.* **23**, 1271–1276 (2017).
5. Chaillon, A. et al. HIV persists throughout deep tissues with repopulation from multiple anatomical sources. *J. Clin. Investig.* **130**, 1699–1712 (2020).
6. Dufour, C. et al. Near full-length HIV sequencing in multiple tissues collected postmortem reveals shared clonal expansions across distinct reservoirs during ART. *Cell Rep.* **42**, 113053 (2023).
7. Sun, W. et al. Persistence of intact HIV-1 proviruses in the brain during antiretroviral therapy. *eLife* **12,** RP89837 (2023).
8. Solis-Leal, A. et al. Lymphoid tissues contribute to plasma viral clonotypes early after antiretroviral therapy interruption in SIV-infected rhesus macaques. *Sci. Transl. Med.* **15**, eadi9867 (2023).
9. Grau-Exposito, J. et al. A novel single-cell FISH-flow assay identifies effector memory CD4(+) T cells as a major niche for HIV-1 transcription in HIV-infected patients. *mBio* **8**, e00876 (2017).
10. Baxter, A. E. et al. Single-cell characterization of viral translation-competent reservoirs in HIV-infected individuals. *Cell host microbe* **20**, 368–380 (2016).
11. Pardons, M. et al. Single-cell characterization and quantification of translation-competent viral reservoirs in treated and untreated HIV infection. *PLoS Pathog.* **15**, e1007619 (2019).
12. Neidleman, J. et al. Phenotypic analysis of the unstimulated in vivo HIV CD4 T cell reservoir. *eLife* **9**, e60933 (2020).
13. Sun, W. et al. Phenotypic signatures of immune selection in HIV-1 reservoir cells. *Nature* **614**, 309–317 (2023).
14. Wu, V. H. et al. Profound phenotypic and epigenetic heterogeneity of the HIV-1-infected CD4(+) T cell reservoir. *Nat. Immunol.* **24**, 359–370 (2023).
15. Wei, Y. et al. Single-cell epigenetic, transcriptional, and protein profiling of latent and active HIV-1 reservoir revealed that IKZF3 promotes HIV-1 persistence. *Immunity* **56**, 2584–2601.e2587 (2023).
16. Prins, J. M. et al. Immuno-activation with anti-CD3 and recombinant human IL-2 in HIV-1-infected patients on potent antiretroviral therapy. *AIDS (Lond., Engl.)* **13**, 2405–2410 (1999).
17. Vemula, S. V. et al. Identification of proximal biomarkers of PKC agonism and evaluation of their role in HIV reactivation. *Antivir. Res.* **139**, 161–170 (2017).
18. Pardons, M. et al. Potent latency reversal by Tat RNA-containing nanoparticle enables multi-omic analysis of the HIV-1 reservoir. *Nat. Commun.* **14**, 8397 (2023).
19. Vibholm, L. K. et al. Characterization of intact proviruses in blood and lymph node from HIV-infected individuals undergoing analytical treatment interruption. *J. Virol.* **93**, e01920–18 (2019).
20. McManus, W. R. et al. HIV-1 in lymph nodes is maintained by cellular proliferation during antiretroviral therapy. *J. Clin. Investig.* **129**, 4629–4642 (2019).
21. Kuo, H. H. et al. Blood and lymph node dissemination of clonal genome-intact human immunodeficiency virus 1 DNA sequences during suppressive antiretroviral therapy. *J. Infect. Dis.* **222**, 655–660 (2020).
22. Recordon-Pinson, P., Gosselin, A., Ancuta, P., Routy, J. P. & Fleury, H. Phylogenetic analysis of HIV-1 archived DNA in blood and gut-associated lymphoid tissue in two patients under antiretroviral therapy. *Gut Pathog.* **13**, 20 (2021).
23. Martin, A. R. et al. Similar frequency and inducibility of intact human immunodeficiency virus-1 proviruses in blood and lymph nodes. *J. Infect. Dis.* **224**, 258–268 (2021).
24. De Scheerder, M. A. et al. HIV rebound is predominantly fueled by genetically identical viral expansions from diverse reservoirs. *Cell host microbe* **26**, 347–358.e347 (2019).
25. Shahid, A. et al. HIV proviral genetic diversity, compartmentalization and inferred dynamics in lung and blood during long-term suppressive antiretroviral therapy. *PLoS Pathog.* **18**, e1010613 (2022).
26. Bozzi, G. et al. No evidence of ongoing HIV replication or compartmentalization in tissues during combination antiretroviral therapy: Implications for HIV eradication. *Sci. Adv.* **5**, eaav2045 (2019).
27. Josefsson, L. et al. The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc. Natl Acad. Sci. USA* **110**, E4987–E4996 (2013).

28. Cole, B. et al. Extensive characterization of HIV-1 reservoirs reveals links to plasma viremia before and during analytical treatment interruption. *Cell Rep.* **39**, 110739 (2022).
29. Banga, R. et al. PD-1(+) and follicular helper T cells are responsible for persistent HIV-1 transcription in treated aviremic individuals. *Nat. Med.* **22**, 754–761 (2016).
30. Pallikkuth, S. et al. Peripheral T follicular helper cells are the major HIV reservoir within central memory CD4 T cells in peripheral blood from chronically HIV-infected individuals on combination antiretroviral therapy. *J. Virol.* **90**, 2718–2728 (2015).
31. Aid, M. et al. Follicular CD4 T helper cells as a major HIV reservoir compartment: a molecular perspective. *Front. Immunol.* **9**, 895 (2018).
32. Ren, Y. et al. BCL-2 antagonism sensitizes cytotoxic T cell-resistant HIV reservoirs to elimination ex vivo. *J. Clin. Investig.* **130**, 2542–2559 (2020).
33. Weymar, G. H. J. et al. Distinct gene expression by expanded clones of quiescent memory CD4(+) T cells harboring intact latent HIV-1 proviruses. *Cell Rep.* **40**, 111311 (2022).
34. Collora, J. A. et al. Single-cell multiomics reveals persistence of HIV-1 in expanded cytotoxic T cell clones. *Immunity* **55**, 1013–1031.e1017 (2022).
35. Lambrechts, L. et al. HIV-PULSE: a long-read sequencing assay for high-throughput near full-length HIV-1 proviral genome characterization. *Nucleic acids Res.* **51**, e102 (2023).
36. Cole, B. et al. In-depth single-cell analysis of translation-competent HIV-1 reservoirs identifies cellular sources of plasma viremia. *Nat. Commun.* **12**, 3727 (2021).
37. Delporte, M. et al. Integrative assessment of total and intact HIV-1 reservoir by a 5-region multiplexed rainbow DNA digital PCR assay. *Clin. Chem.* **71**, 203–214 (2025).
38. Van Gassen, S. et al. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytom. Part A: J. Int. Soc. Anal. Cytol.* **87**, 636–645 (2015).
39. Kearney, M. F. et al. Lack of detectable HIV-1 molecular evolution during suppressive antiretroviral therapy. *PLoS Pathog.* **10**, e1004010 (2014).
40. McMyn, N. F. et al. The latent reservoir of inducible, infectious HIV-1 does not decrease despite decades of antiretroviral therapy. *J Clin. Invest.* **133,** e171554 (2023).
41. Gantner, P. et al. HIV rapidly targets a diverse pool of CD4(+) T cells to establish productive and latent infections. *Immunity* **56**, 653–668.e655 (2023).
42. Sallusto, F., Geginat, J. & Lanzavecchia, A. Central memory and effector memory T cell subsets: function, generation, and maintenance. *Annu. Rev. Immunol.* **22**, 745–763 (2004).
43. Crotty, S. T follicular helper cell biology: a decade of discovery and diseases. *Immunity* **50**, 1132–1148 (2019).
44. Locci, M. et al. Human circulating PD-1+CXCR3-CXCR5+ memory Tfh cells are highly functional and correlate with broadly neutralizing HIV antibody responses. *Immunity* **39**, 758–769 (2013).
45. Morita, R. et al. Human blood CXCR5(+)CD4(+) T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion. *Immunity* **34**, 108–121 (2011).
46. Cenerenti, M., Saillard, M., Romero, P. & Jandus, C. The Era of Cytotoxic CD4 T Cells. *Front. Immunol.* **13**, 867189 (2022).
47. Appay, V. et al. Characterization of CD4(+) CTLs ex vivo. *J. Immunol. (Baltim., Md.: 1950)* **168**, 5954–5958 (2002).
48. Buggert, M. et al. The identity of human tissue-emigrant CD8(+) T Cells. *Cell* **183**, 1946–1961.e1915 (2020).
49. Appay, V. et al. HIV-specific CD8(+) T cells produce antiviral cytokines but are impaired in cytolytic function. *J. Exp. Med.* **192**, 63–75 (2000).
50. Gordon, C. L. et al. Tissue reservoirs of antiviral T cell immunity in persistent human CMV infection. *J. Exp. Med.* **214**, 651–667 (2017).
51. Shacklett, B. L. et al. Trafficking of human immunodeficiency virus type 1-specific CD8+ T cells to gut-associated lymphoid tissue during chronic infection. *J. Virol.* **77**, 5621–5631 (2003).
52. Altfeld, M. et al. Expansion of pre-existing, lymph node-localized CD8+ T cells during supervised treatment interruptions in chronic HIV-1 infection. *J. Clin. Investig.* **109**, 837–843 (2002).
53. Morcilla, V. et al. HIV-1 Genomes Are Enriched in Memory CD4(+) T-Cells with Short Half-Lives. *mBio* **12**, e0244721 (2021).
54. Bacchus-Souffan, C. et al. Relationship between CD4 T cell turnover, cellular differentiation and HIV persistence during ART. *PLoS Pathog.* **17**, e1009214 (2021).
55. Gantner, P. et al. Single-cell TCR sequencing reveals phenotypically diverse clonally expanded cells harboring inducible HIV proviruses during ART. *Nat. Commun.* **11**, 4089 (2020).
56. Wang, Z. et al. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proc. Natl Acad. Sci. USA* **115**, E2575–E2584 (2018).
57. Yokoi, T. et al. Identification of a unique subset of tissue-resident memory CD4(+) T cells in Crohn's disease. *Proc. Natl Acad. Sci. USA* **120**, e2204269120 (2023).
58. Kunzli, M. & Masopust, D. CD4(+) T cell memory. *Nat. Immunol.* **24**, 903–914 (2023).
59. Richard, A. C. Divide and conquer: phenotypic and temporal heterogeneity within CD8(+) T cell responses. *Front. Immunol.* **13**, 949423 (2022).
60. Zarate, S., Pond, S. L., Shapshak, P. & Frost, S. D. Comparative study of methods for detecting sequence compartmentalization in human immunodeficiency virus type 1. *J. Virol.* **81**, 6643–6651 (2007).
61. Lee, G. Q. Chemistry and bioinformatics considerations in using next-generation sequencing technologies to inferring HIV proviral DNA genome-intactness. *Viruses* **13**, 1874 (2021).
62. White, J. A. et al. Measuring the latent reservoir for HIV-1: Quantification bias in near full-length genome sequencing methods. *PLoS Pathog.* **18**, e1010845 (2022).
63. Vaser, R., Sovic, I., Nagarajan, N. & Sikic, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* **27**, 737–746 (2017).
64. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
65. Wright, I. A. et al. HIVIntact: a python-based tool for HIV-1 genome intactness inference. *Retrovirology* **18**, 16 (2021).
66. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059–3066 (2002).
67. Jensen, M. A. et al. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J. Virol.* **77**, 13376–13388 (2003).
68. Lengauer, T., Sander, O., Sierra, S., Thielen, A. & Kaiser, R. Bioinformatics prediction of HIV coreceptor usage. *Nat. Biotechnol.* **25**, 1407–1410 (2007).
69. Deng, W. et al. DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. *Biotechniques* **48**, 405–408 (2010).
70. Kosakovsky Pond, S. L. et al. HyPhy 2.5-a customizable platform for evolutionary hypothesis testing using phylogenies. *Mol. Biol. evolution* **37**, 295–299 (2020).
71. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
72. Lambrechts, L. et al. HIV-PULSE: a long-read sequencing assay for high-throughput near full-length HIV-1 proviral genome characterization. *Github* **51**, e102 (2023).

73. Cole, B. et al. In-depth single-cell analysis of translation-competent HIV-1 reservoirs identifies cellular sources of plasma viremia. *Github* **12**, 3727 (2021).

Acknowledgements

We thank all participants who donated blood and tissue samples, as well as MDs and study nurses who helped with the recruitment and coordination of this study. The study team thanks Maarten Verdonckt and Sophie Vermaut for assisting with the flow cytometry sorting experiments. We thank Jinho Park from Arcturus Therapeutics for formulating the LNP. This current research work was supported by VLAIO O&O (HBC.2018.2278), by NIH/NIAID (1 UM1 AI164561-01, NIH MDC grant: RID-HIV UM1AI164561) and by the Research Foundation Flanders (G0B3820N/G0A1824N). MP was supported by postdoctoral funding from VLAIO O&O (HBC.2018.2278). LL was supported by FWO Vlaanderen (1S29220N). SG was supported by FWO Vlaanderen (G0A1824N). LV was supported by the Research Foundation Flanders (1.8.020.09.N.00) and the Collen-Francqui Research Professor Mandate. The content is entirely the responsibility of the authors and may not reflect the official views of the National Institutes of Health.

Author contributions

M.P., L.L., S.G., and L.V. conceptualized the experiments. M.P., L.L., Y.N., L.T., S.D.B. performed experiments. L.L. analyzed viral sequencing data and constructed phylogenetic trees. E.V.G. helped with the coordination of the study. JV contributed to the design and formulation of the lipid nanoparticle. M.P. and L.L. wrote the paper. All authors read and edited the paper.

Competing interests

The Authors M.P., L.L., Y.N., L.T., S.D.B., S.G. and L.V. declare no competing interests. The Author E.V.G. declares the following competing interests: this author was an employee of Johnson & Johnson and may be Johnson & Johnson stockholder. The Author J.V. declares the following competing interests: this author was an employee of Arcturus Therapeutics and may be Arcturus Therapeutics stockholder. Johnson & Johnson has a pending patent application related to the Tat-LNP mRNA molecule.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-57332-5.

Correspondence and requests for materials should be addressed to Linos Vandekerckhove.

Peer review information *Nature Communications* thanks Andrew Atkins, Heng-Chang Chen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.